Proc. R. Soc. Lond. B **235**, 221–245 (1988)
Printed in Great Britain

Feature detection in human vision: a phase-dependent energy model

BY M. CONCETTA MORRONE† AND D. C. BURR†

Department of Psychology, University of Western Australia, Nedlands, Western Australia 6009, *Australia*

(*Communicated by H. B. Barlow, F.R.S.* – *Received* 22 *February* 1988 – *Revised* 20 *June* 1988)

[Plates 1–4]

This paper presents a simple and biologically plausible model of how mammalian visual systems could detect and identify features in an image. We suggest that the points in a waveform that have unique perceptual significance as 'lines' and 'edges' are the points where the Fourier components of the waveform come into phase with each other. At these points 'local energy' is maximal. Local energy is defined as the square root of the sum of the squared response of sets of matched filters, of identical amplitude spectrum but differing in phase spectrum by 90°: one filter type has an even-symmetric line-spread function, the other an odd-symmetric line-spread function. For a line the main contribution to the local energy peak is in the output of the even-symmetric filters, whereas for edges it is in the output of the odd-symmetric filters. If both filter types respond at the peak of local energy, both edges and lines are seen, either simultaneously or alternating in time. The model was tested with a series of images, and shown to predict well the position of perceived features and the organization of the images.

INTRODUCTION

To function efficiently within reasonable informational limits, any visual system (biological or artificial) must simplify the image and record it in an economical 'token' form (see, for example, Barlow 1959; Marr 1976). Most researchers agree that edges and lines are particularly rich sources of image information, and that organization of these features could provide the basis for an efficient description of the image. Several models of edge and line detection have been developed.

An early idea was that there may exist in the human visual system detectors specialized to respond to edges and lines. Tolhurst (1972) and Kulikowski & King-Smith (1973) suggested that visual detectors may take advantage of the local symmetry of these features: a line is an even-symmetric function ($f(x) = f(-x)$, choosing the line centre as origin), and an edge an odd-symmetric function ($f(x) = -f(-x)$, choosing the mean luminance crossing as origin). Detectors with even-symmetric line-spread functions, or *receptive fields*, will respond best when

† Present address: Istituto di Neurofisiologia del C.N.R., via S. Zeno 51, Pisa 56100, Italy.

centred on a line; those with odd-symmetric fields respond best when centred on an edge.

Psychophysical evidence is consistent with the suggestion that visual detectors have even- and odd-symmetric receptive fields (see, for example, Tolhurst 1972; Shapley & Tolhurst 1973; Kulikowski & King-Smith 1973; Field & Nachmias 1984; Burr *et al.* 1988; but see also Field & Tolhurst 1986). However, the mechanism by which the hypothetical detectors may locate edges and lines is not obvious. The problem is that an edge-detector (odd-symmetric field) will respond both to an edge centred on the receptive field and to a line positioned slightly away from the field centre. The inappropriate response to a line can be almost as great as the appropriate response to an edge. The same argument applies to line-detectors, with even-symmetric fields (examples of output from even- and odd-symmetric operators are shown in figure 3). Thus the two detectors do not give an unambiguous response, but one which must be decoded by additional rules or restraints. This limitation is particularly evident with algorithms designed for robotic vision. For example, Canny (1983) used two operators of even and odd symmetry and searched for local maxima. However, the same feature was marked by both operators, usually at different points.

Some models of human feature detection employ only one operator. For example, Marr & Hildreth (1980) use as operators the Laplacian of a Gaussian function with different space constants, and search the output for coincidence of zero-crossings across scales. This model is designed to detect edges, not lines. The requirement for coincidence of zero-crossings across scales eliminates the inappropriate marking of lines, which occur at different positions at different scales. However, this restriction can also eliminate real edges under certain conditions. One example is when two edges occur nearby: zero-crossings at larger scales will occur midway between the edges. Other examples where the model fails will be presented later.

These are several problems inherent in most existing models of feature detection. A linear filter operation, the first stage of most models, creates 'ringing', which can create spurious peaks and zero-crossings. These can be difficult to distinguish from peaks or zero-crossings associated with real features. Low amounts of visual noise can also cause spurious responses, again difficult to remove. Besides the technical problems outlined above, a deeper theoretical limitation is that there exists no satisfactory mathematical definition of what constitutes for the visual system a line or an edge.

In this paper we propose a new model of feature detection, based on a new definition of lines and edges. The definition is best understood by considering the local Fourier representation of edges and lines. In Fourier space, the symmetry of edges and lines is reflected in the phase spectrum. An isolated line (even-symmetric function) expands to a series of cosine components (choosing the line centre as origin), with all sine components at zero amplitude. An isolated edge (odd-symmetric function) expands to a series of sine components (choosing the mean luminance crossing point as origin), with all cosine components at zero amplitude. Thus when the edge or line is chosen as origin, the Fourier phase spectrum is constant: zero at all frequencies for a line and $\pi/2$ (or 90°) for an edge. For other choices of origin the Fourier phase spectrum will not be constant. However, for

any origin, at the point of an isolated line or edge the *arguments* (which can also be termed *arrival phases*) of all Fourier harmonics will be identical.

This fact suggests a potentially useful property of lines and edges: they occur at points of the waveform where the arrival phases of the Fourier components are maximally similar. The value of the average arrival phase at that point determines the nature of the feature: values near zero correspond to a line, and values near $\pi/2$ correspond to an edge. In previous publications (Morrone *et al.* 1986; Ross *et al.* 1988) we have shown that this definition predicts successfully 'Mach bands', the illusory light and dark stripes seen where luminance gradients meet plateaux (Mach 1865). At the points where the bands are seen, the arrival phases of the Fourier components are maximally similar, and near zero, even though there is no line-like discontinuity in the luminance profile at the point where the bands are seen. That the definition applies to illusory as well as real features suggests that it may have general validity.

The present study has two goals: to test whether 'congruence of arrival phase' may be a useful and general definition of visual features; and to devise a biologically plausible model of feature detection based on this definition. The following section introduces the model, shows mathematically how it relates to the phase definition of features, and gives the details of the parameters chosen for implementation. The remainder of the paper reports a series of demonstrations and experiments that illustrate the relation between arrival-phase congruence and the perception of visual features. The images used in the study were created by manipulating the phase spectrum of a simple one-dimensional waveform, the squarewave. For each image, we test the performance of the model and show that it predicts successfully the position and nature of the main features.

Model

The basic operators of the model are pairs of filters of equal amplitude spectra but orthogonal in phase: one has a Fourier phase spectrum of 0 (cosine phase), the other a Fourier phase spectrum of $\pi/2$ (sine phase). The line-spread functions of the filters are symmetric, one with even symmetry (F_e) the other with odd symmetry (F_o). F_o is the Hilbert transform of F_e (see, for example, Garnett 1981). Provided that the functions are contained in L^2 space (functions for which the integral of their square is finite) they will be orthogonal. That is

$$\int F_e(x)\, F_o(x)\, \mathrm{d}x = 0. \tag{1}$$

The same is true for the response of the two filters to any image $I(x)$.

Given that

$$\left.\begin{aligned} O_e(x) &= \int F_e(x-\zeta)\, I(\zeta)\, \mathrm{d}\zeta, \\ O_o(x) &= \int F_o(x-\zeta)\, I(\zeta)\, \mathrm{d}\zeta, \end{aligned}\right\} \tag{2}$$

it follows that

$$\int O_e(x)\, O_o(x)\, \mathrm{d}x = 0. \tag{3}$$

The proof of orthogonality follows from the fact that the even and odd line-spread functions are the Hilbert transforms of each other (see, for example, Garnett 1981) and that the Hilbert transform is a skew-symmetric operator.

In the space spanned by the orthogonal functions O_e and O_o the response of each filter is represented by a curve parametric in x, given by equation (2) (see figure 1). By considering a polar coordinate representation, the same information can be described by the function's modulus E and argument A given by:

$$\left.\begin{aligned} E(x) &= \sqrt{[O_o^2(x)+O_e^2(x)]}, \\ A(x) &= \arctan\,[O_o(x)/O_e(x)]. \end{aligned}\right\} \tag{4}$$

We term the function $E(x)$ the *local energy function* (following Adelson & Bergen 1985): it represents the vector length of the combined response at the point x.

Figure 1 illustrates how the model relates to the phase of the Fourier components of an image. A periodic function $I(x)$ can be expanded as a series of cosine harmonic components of amplitude a_n and phase ϕ_n. That is,

$$I(x) = \sum_{n=1}^{\infty} a_n \cos\,(nx\omega+\phi_n), \tag{5}$$

where ω is the fundamental frequency.

The output functions O_e and O_o of the filters to the input function $I(x)$ will be

$$\left.\begin{aligned} O_e &= \sum_{n=1}^{\infty} b_n \cos\,(nx\omega+\phi_n), \\ O_o &= \sum_{n=1}^{\infty} -b_n \sin\,(nx\omega+\phi_n), \end{aligned}\right\} \tag{6}$$

where b_n is the amplitude of each harmonic of the output functions, given by the product of the input amplitudes a_n and the filter gain at frequency $n\omega$. Note that the even-symmetric filters do not alter the input phase ϕ_n, whereas the odd-symmetric filters cause a phase shift of $\pi/2$.

In the space spanned by the functions O_e and O_o, the nth harmonic component of the filtered output is represented by a vector of length b_n and argument $A_n(x) = (nx\omega+\phi_n)$. For each spatial position x, the harmonic components sum vectorially to a resultant vector of length $E(x)$ (local energy) and argument $A(x)$. Because the amplitudes b_n are independent of the parameter x, local maxima of the energy function $E(x)$ occur at the points where the argument $A_n(x)$ of the harmonics is most similar. The argument $A_n(x)$ can also be referred to as the *arrival phase* of the harmonic at that point (not to be confused with ϕ_n, the Fourier phase of the harmonic, which depends on the choice of origin). Thus local maxima of the energy function $E(x)$ occur at points where the arrival phases are most similar. At the points of maximum local energy, the average arrival phase is given by the argument $A(x)$. In the example of figure 1, the argument is $\pi/2$ and $3\pi/2$ at the peaks, indicating a positive- and negative-going edge.

We suggest that the visual system could locate features of interest by searching

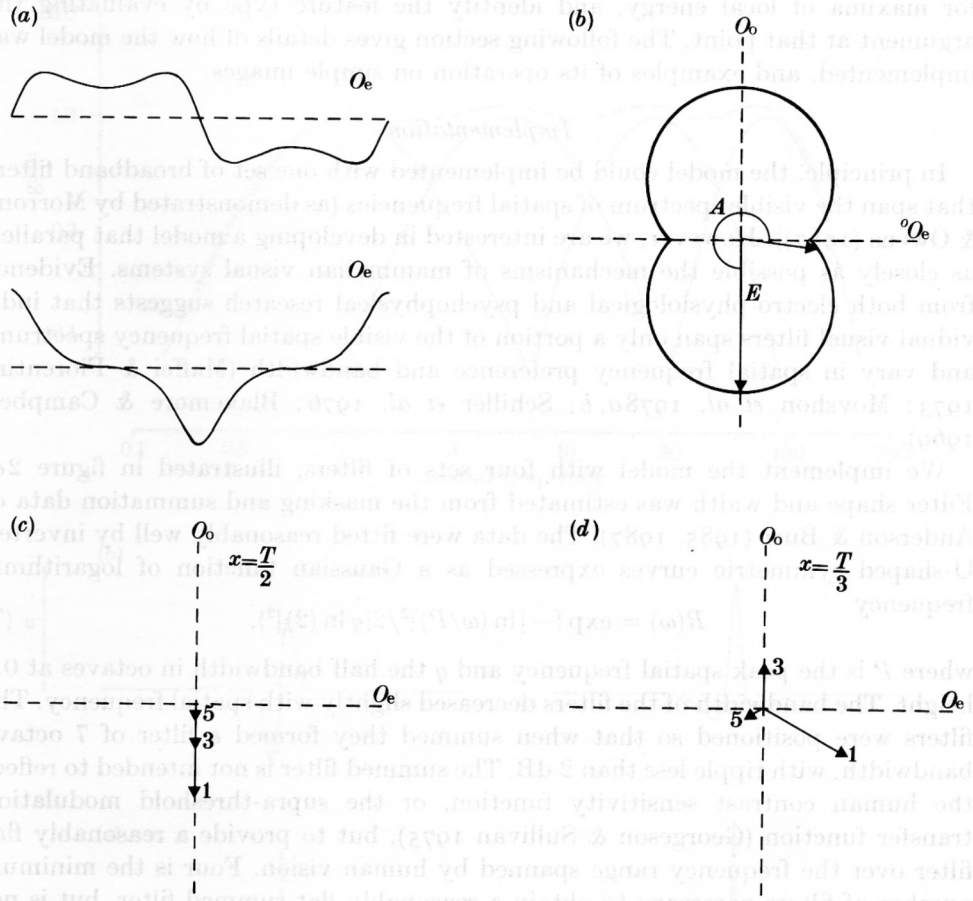


FIGURE 1. Illustration of the local energy function, and how it relates to Fourier phase. (*a*) The curves show the responses of a pair of matched filters to a squarewave. The filters are centred at 2 cycles per period with half-bandwidth of 1 octave, and are in quadrature phase. Their responses (O_e and O_o) are given by

$$\left.\begin{aligned} O_e(x) &= 4a/\pi \sum_{k=1}^{\infty} 1/k \cos(2\pi kx/T - \pi/2) \exp\{-[\ln(k/P)]^2/2[q \ln(2)]^2\}, \\ O_o(x) &= 4a/\pi \sum_{k=1}^{\infty} -1/k \sin(2\pi kx/T - \pi/2) \exp\{-[\ln(k/P)]^2/2[q \ln(2)]^2\}, \end{aligned}\right\} \quad \text{(F 1)}$$

for k odd integer, where T is the period of the squarewave, a is its amplitude, P the peak spatial frequency of the filters and q their bandwidth (in octaves). As the functions are orthogonal, the same information can be represented by plotting O_o against O_e for each value of x. (*b*) The plot is characterized by its polar coordinates E (local energy) and argument A. Note that E has maximum length when $A = \pm\pi/2$. (*c*, *d*) The contribution of single Fourier components of the filtered squarewave for two values of x. For $x = T/2$, the arguments A_n of all components are $\pi/2$. The vectorial sum of the components gives the maximum value of E ($1.3a$). For $x = T/3$, the arguments A_n vary from one component to another, so the vectorial sum will give a reduced value of E. In this case, the vectors sum to produce local energy E of $0.6a$, and argument A of 1.96. E will be minimal for $x = T/4$, where the first and third harmonics oppose each other.

for maxima of local energy, and identify the feature type by evaluating the argument at that point. The following section gives details of how the model was implemented, and examples of its operation on simple images.

Implementation

In principle, the model could be implemented with one set of broadband filters that span the visible spectrum of spatial frequencies (as demonstrated by Morrone & Owens (1987)). However, we are interested in developing a model that parallels as closely as possible the mechanisms of mammalian visual systems. Evidence from both electro-physiological and psychophysical research suggests that individual visual filters span only a portion of the visible spatial frequency spectrum, and vary in spatial frequency preference and bandwidth (Maffei & Fiorentini 1973; Movshon *et al.* 1978*a*, *b*; Schiller *et al.* 1976; Blakemore & Campbell 1969).

We implement the model with four sets of filters, illustrated in figure 2*a*. Filter shape and width was estimated from the masking and summation data of Anderson & Burr (1985, 1987). The data were fitted reasonably well by inverted U-shaped symmetric curves expressed as a Gaussian function of logarithmic frequency

$$R(\omega) = \exp\{-[\ln(\omega/P)]^2/2[q\ln(2)]^2\}, \tag{7}$$

where P is the peak spatial frequency and q the half bandwidth in octaves at 0.6 height. The bandwidth of the filters decreased slightly with spatial frequency. The filters were positioned so that when summed they formed a filter of 7 octave bandwidth, with ripple less than 2 dB. The summed filter is not intended to reflect the human contrast sensitivity function, or the supra-threshold modulation transfer function (Georgeson & Sullivan 1975), but to provide a reasonably flat filter over the frequency range spanned by human vision. Four is the minimum number of filters necessary to obtain a reasonably flat summed filter, but is not intended to reflect the number that may actually exist in human vision (cf. Wilson *et al.* 1983). In fact, the exact shape, bandwidth and number of filters are not at all crucial for the performance of the model (see Discussion).

On the basis of psychophysical evidence (Field & Nachmias 1984; Burr *et al.* 1988), the phase spectrum of the filters was assumed to be constant, either at 0 or $\pi/2$. Inverse Fourier transformation of the tuning functions gives the receptive field profiles shown in figures 2*b* and 2*c* for two of the tuning curves. Two differences are evident between the receptive fields of high and low frequency preference. Those preferring high spatial frequency have more ripples (resulting from the narrow tuning), and higher peak activity. The receptive fields shown here are similar to those of simple cells in cat and monkey primary visual cortex (Hubel & Wiesel 1962, 1977; Kulikowski & Bishop 1981), and are also consistent with the receptive field size and shape inferred from psychophysical summation studies (Anderson & Burr 1987).

Figure 3 gives examples of the implementation of the model for three input waveforms: a squarewave (3*a*), a series of lines (3*b*) and a trianglewave (3*c*). Each input waveform was filtered by the four pairs of matched filters, to produce eight output waveforms (figure 3*g*–*i*). At each scale the response of the matched filters

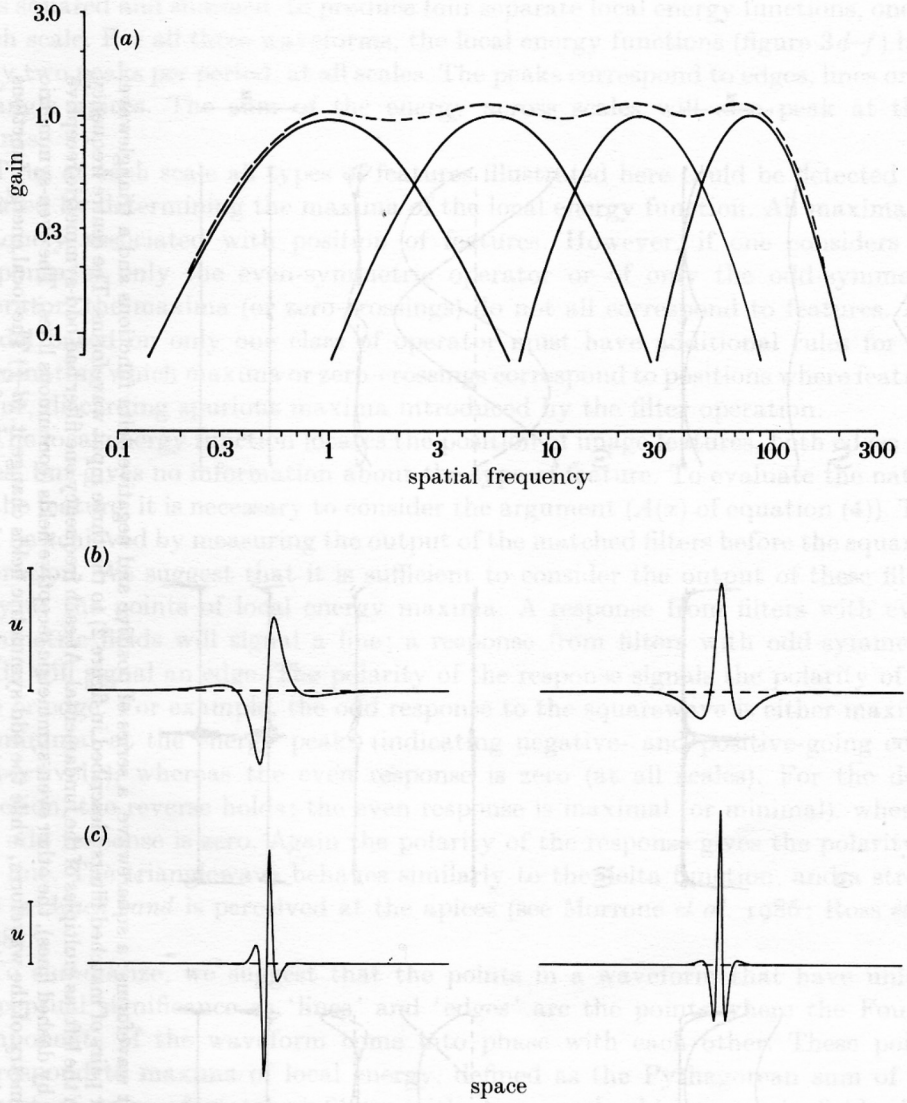


FIGURE 2. (*a*) Spatial frequency tuning (gain) of the four sets of filters used in the simulations. Their frequency response $R(\omega)$ is given by

$$R(\omega) = \exp\{-[\ln(\omega/P)]^2/2[q\ln(2)]^2\}, \qquad \text{(F 2)}$$

where P is the peak spatial frequency and q the half bandwidth in octaves at 0.6 height (equation (7) in text). The centre frequencies P (in cycles per period of input waveform) and bandwidths q (in octaves) were, respectively: 1, 1.2; 5.7, 0.9; 27.4, 0.73; 65.2, 0.57. Bandwidth decreased slightly with increasing spatial frequency, in agreement with psychophysical estimates. (*b*, *c*) Impulse response functions of the second lowest (*b*) and second highest filter (*c*). The functions on the left (odd-symmetric) assume a constant phase spectrum of $\pi/2$ and those on the right (even-symmetric) a constant phase spectrum of 0. The calibration u on the left represents one unit of amplitude.

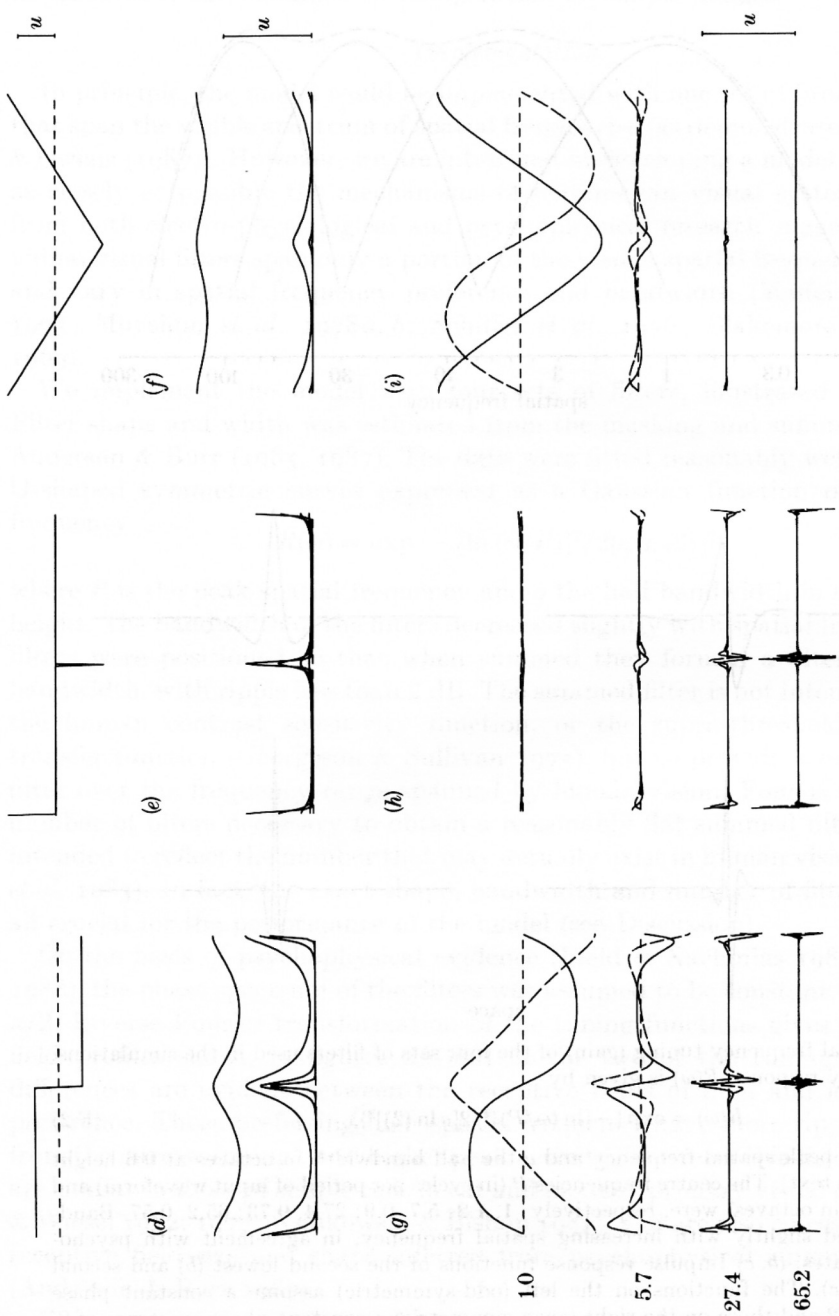


FIGURE 3. (*a–c*) Three input waveforms: a squarewave, a series of positive and negative delta functions and a trianglewave. (*g–i*) The output of four pairs of matched filters (described in figure 2) to the three input waveforms. The centre frequency of the filters is shown on the left, as multiples of the fundamental. At each frequency one filter had odd-symmetric receptive fields (output indicated by dashed lines), the other even-symmetric receptive fields (continuous lines). The even-symmetric fields preserve the symmetry of the waveform, while the odd-symmetric fields change it. (*d–f*) The local energy functions at four scales, calculated by taking the square root of the sum of the squared responses of the even- and odd-symmetric receptive fields. The more sharply peaked profiles correspond to smaller scales. For all the functions, at all scales, the local energy functions peak at the feature of interest for the visual system: the edge, the line or the Mach band. These are the only peaks in the energy function: there are no other ripples or ringing (compare with the filter outputs of (*g–i*)).

was squared and summed, to produce four separate local energy functions, one for each scale. For all three waveforms, the local energy functions (figure 3*d–f*) have only two peaks per period, at all scales. The peaks correspond to edges, lines or the triangle apices. The sum of the energy across scales will also peak at these points.

Thus at each scale all types of features illustrated here could be detected and located by determining the maxima of the local energy function. All maxima are uniquely associated with position of features. However, if one considers the response of only the even-symmetric operator or of only the odd-symmetric operator, the maxima (or zero-crossings) do not all correspond to features. Any model based on only one class of operator must have additional rules for discriminating which maxima or zero-crossings correspond to positions where features occur, discarding spurious maxima introduced by the filter operation.

The local energy function locates the position of image features, both edges and lines, but gives no information about the type of feature. To evaluate the nature of the feature, it is necessary to consider the argument ($A(x)$ of equation (4)). This can be achieved by measuring the output of the matched filters before the squaring operation. We suggest that it is sufficient to consider the output of these filters only at the points of local energy maxima. A response from filters with even-symmetric fields will signal a line; a response from filters with odd-symmetric fields will signal an edge. The polarity of the response signals the polarity of the line or edge. For example, the odd response to the squarewave is either maximal or minimal at the energy peaks (indicating negative- and positive-going edges respectively), whereas the even response is zero (at all scales). For the delta function, the reverse holds: the even response is maximal (or minimal), whereas the odd response is zero. Again the polarity of the response gives the polarity of the line. The trianglewave behaves similarly to the delta function, and a strong line or *Mach band* is perceived at the apices (see Morrone *et al.* 1986; Ross *et al.* 1988).

To summarize, we suggest that the points in a waveform that have unique perceptual significance as 'lines' and 'edges' are the points where the Fourier components of the waveform come into phase with each other. These points correspond to maxima of local energy, defined as the Pythagorean sum of the output of pairs of matched filters with even- and odd-symmetric fields. The following sections test these assertions with a series of demonstrations and experiments with images derived from simple one-dimensional patterns.

Methods

The stimuli used in this study were generated by computer (NCR Tower) from the formulae given in the appropriate sections, and displayed on a black and white video screen (Sony, PVM-91CE). The resolution was 512×512 pixels, with 256 grey levels per pixel. To avoid aliasing and ringing effects (such as the Gibbs phenomenon, see Hewitt & Hewitt (1979)), Fourier components above the Nyquist frequency limit were rejected, and the amplitude spectra multiplied by a shallow Gaussian $G(k)$

$$G(k) = \exp(-k^2/2\sigma^2) \tag{8}$$

where k is an integer multiple of the fundamental harmonic frequency. For all stimuli except figure 6, σ was half the Nyquist frequency. For figure 6, σ was 2 cycles per period. All computations were done with a modified version of the HIPS image-processing system (Landy *et al.* 1984*a*, *b*).

Waveforms were displayed on the video screen, and photographed. Observers were invited to view both the screen and the photographs and describe their appearance. Many observers were initially unaware of the aims of the experiment, and of the profile of the patterns they were viewing.

For experiment 1, observers were required to indicate the apparent position of perceived lines and edges on a waveform. The video monitor was optically superimposed on an oscilloscope which displayed a small dot. Observers could adjust the position of the dot to match the apparent position of lines and edges, by turning a potentiometer. A small computer read the position of the potentiometer and moved the dot horizontally. The stimulus occupied 10.5 cm × 10.5 cm on the monitor (subtending 2° at 3 m viewing distance) and had mean luminance of 250 cd m^{-2}.

Results and simulations

This section examines the perceptual effects of manipulation of the phase and amplitude spectra of squarewaves, and compares these with predictions of the proposed model. The Fourier expansion of a squarewave (choosing as origin the negative to positive transition) can be expressed as

$$S(x) = L_0 + 4a/\pi \sum_{k=1}^{\infty} 1/k \cos(2\pi kx/T - \pi/2), \tag{9}$$

where k is an odd integer, L_0 is the mean luminance, T is the period, and a is the mean-to-peak amplitude of the squarewave, and also the standard deviation of the luminance of the waveform, which does not change with manipulation of the phase spectrum (from Parseval's theorem). For this reason, we define contrast of all waveforms as the ratio of standard deviation to mean (a/L_0). For squarewaves, this ratio is equivalent to Michelson contrast.

Inspection of equation (9) shows that the argument, or arrival phases, of all harmonics (given by $2\pi kx/T - \pi/2$) are identical at two points per period: at the start of each period ($x = iT$, for integer i) they are all $-\pi/2$, and at the centre of each period ($x = T(i+1/2)$) they are $\pi/2$. These points occur at the edges of the squarewave.

Demonstration 1

The purpose of this experiment was to demonstrate the relation between congruence of arrival phase and the appearance of visual features, and to test the performance of the model on these images. The waveforms of figure 4, plate 1, all have the amplitude spectrum of a squarewave, but differ in phase. For each waveform, the Fourier phases of all cosine harmonics were ϕ at origin. The general equation for the waveforms is

$$f(x) = L_0 + 4a/\pi \sum_{k=1}^{\infty} G(k)/k \cos(2\pi kx/T - \phi). \tag{10}$$

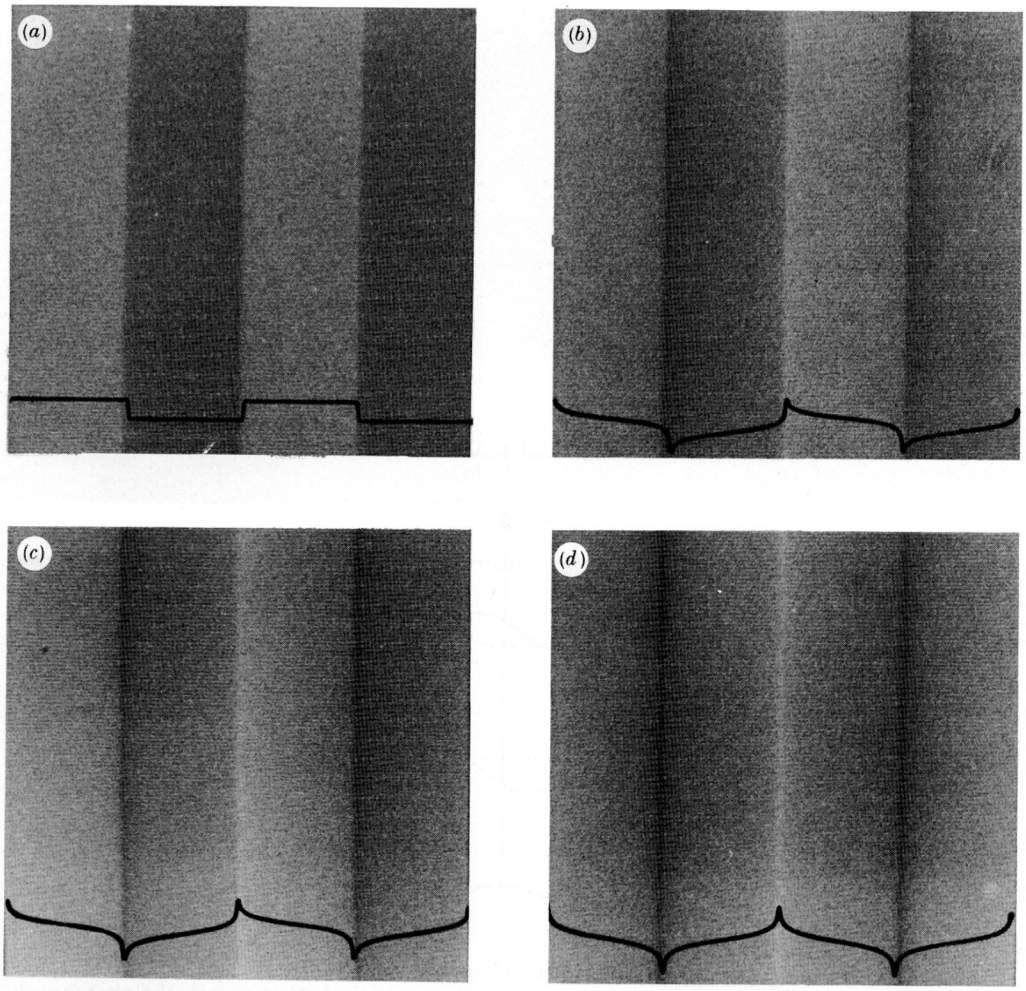


FIGURE 4. Reproductions of two periods of a grating with luminance profile $f(x)$ defined by

$$f(x) = L_0 + 4a/\pi \sum_{k=1}^{\infty} G(k)/k \cos(2\pi kx/T - \phi), \qquad \text{(F 3)}$$

for k odd integer and $0 < x < 2T$. L_0 is mean luminance, a amplitude, T period and ϕ phase (equation (10) in text). The origin is at the left-most point of each pattern. Phase ϕ was $\pi/2$, $\pi/3$, $\pi/4$ and 0 (90°, 60°, 45° and 0°) for (*a*), (*b*), (*c*) and (*d*) respectively. For all figures, the major features seem to occur in the same position, despite the differences in luminance profiles. The features occur at the only points on the waveform where the arguments (or arrival phases) of all Fourier components phases are identical: at $x = 0, T/2, T, 3T/2$ and T. The nature of the feature changes progressively from edge to bar, with decreasing ϕ.

(*Facing p.* 230)

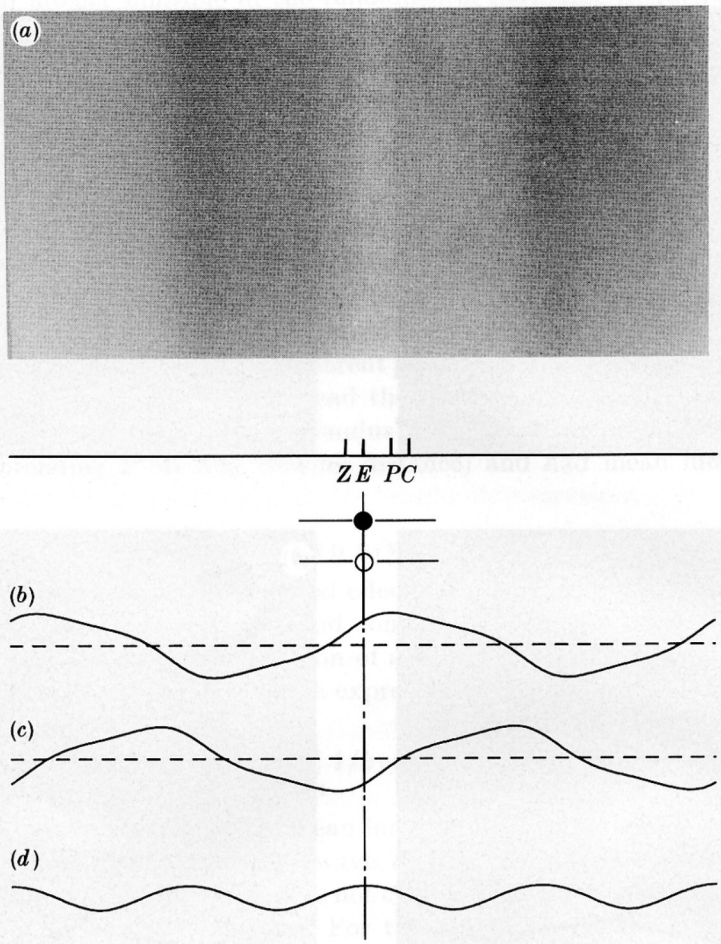


FIGURE 6. (*a*) Waveform with constant Fourier phase spectrum of $\pi/4$, filtered with the low-pass Gaussian filter of equation (8), with $\sigma = 2$ cycles per period. Underneath the photograph are the original waveform, its Hilbert transform and the local energy profile. Also shown are the positions of maximum luminance slope (*Z*), peak in local energy (*E*), peak in luminance (*P*) and the centroid (*C*) of the mean-luminance bounded mass (from left to right). After observing the photograph for a while, its appearance seems to alternate from a squarewave configuration to a series of lines. Observers were asked to indicate separately the apparent position of the edges of the squarewave and the centre of the lines. The means of ten settings for two observers are shown by the open circle (for lines) and closed circle (for edges). Both apparent positions are closer to the local peak in energy than any of the other predictions. The reader can readily verify this result.

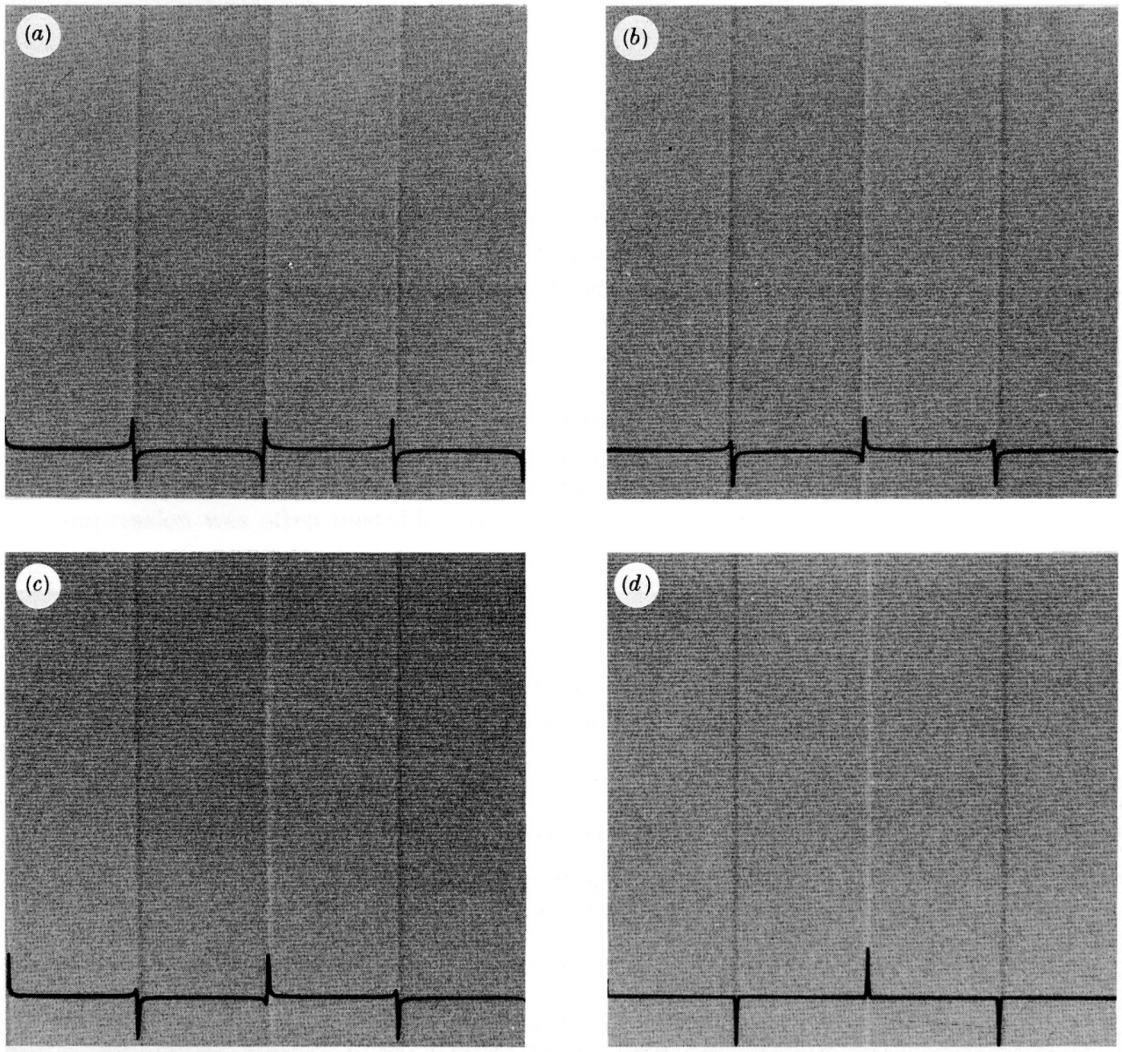
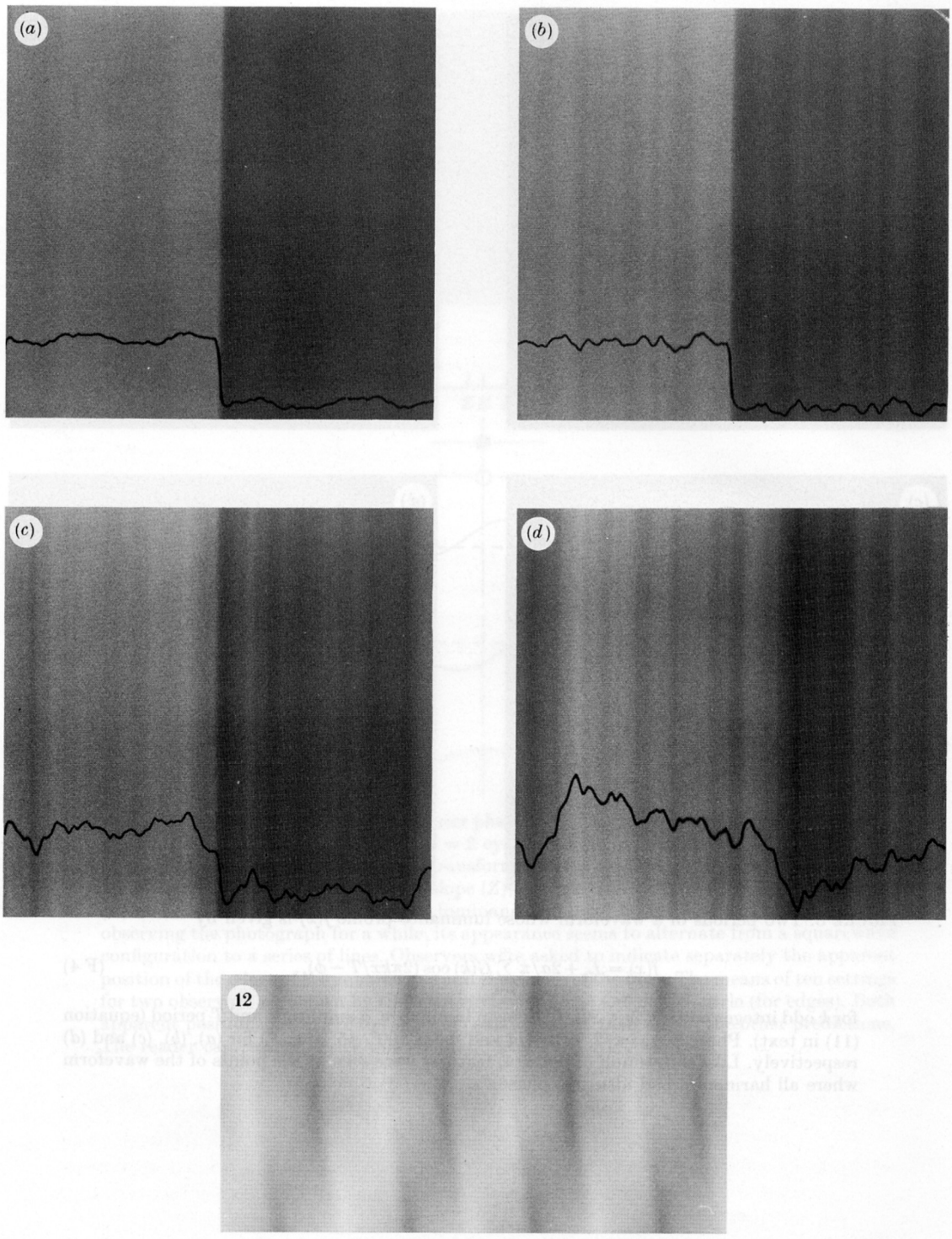


FIGURE 8. Two periods of a waveform whose luminance profile $f(x)$ is given by

$$f(x) = L_0 + 2a/\pi \sum_{k=1}^{\infty} G(k) \cos(2\pi kx/T - \phi), \qquad \text{(F 4)}$$

for k odd integer and $0 < x < 2T$. L_0 is mean luminance, a amplitude and T period (equation (11) in text). Phase ϕ was $\pi/2$, $\pi/3$, $\pi/4$ and 0 (90°, 60°, 45° and 0°) for (*a*), (*b*), (*c*) and (*d*) respectively. Like the stimuli of figure 4, features were seen at the points of the waveform where all harmonics had identical phase.



FIGURES 10 AND 12. For description see opposite.

Where k is an odd integer, a is the standard deviation of the luminance values of the waveform and $G(k)$ is the smoothing Gaussian of equation (8). The value of ϕ determines the phase at origin.

Figure 4 reproduces two periods of the patterns at about 6% contrast (a/L_0), with their luminance profiles. Observers were asked to indicate the major features of the patterns (presented without luminance profiles), and describe the nature of the features. Figure 4*a* (the squarewave) was seen as four broad stripes separated by three clear edges, two negative-going and one positive-going. The edges are considered to be the features of importance, separating regions of homogeneous luminance. Figure 4*d* ($\phi = 0°$) also had three features, but the features appeared to be (slightly blurred) lines, two dark and one bright. The regions between the lines seemed fairly homogeneous. For figures 4*b* and 4*c* ($\phi = 60°$ and 45°), observers reported three edges with superimposed lines, somewhat like a Mach band. Some (but not all) observers claimed that the patterns were unstable, with the relative salience of edges and lines fluctuating rapidly over time. Most observers reported an impression of depth in figure 4*b–d*, with corrugations like a radiator. Again, this impression was often unstable, giving way to a flat interpretation of edges and lines.

It is interesting that varying ϕ did not alter the apparent position of the features. Varying the phase by ϕ shifts all harmonics by an amount inversely proportional to spatial frequency. For the fundamental the shift is $\phi T/2\pi$, which equals $T/4$ in the extreme case. However, as the points where all harmonics come into phase are the same for all patterns, the features always occur at the same position. What changes is the type of feature.

Description of plate 4

Figure 10. One cycle of gratings derived from squarewaves, with phase randomized over a given range. The equation for their luminance profile is

$$f(x) = L_0 + 4a/\pi \sum_{k=1}^{\infty} G(k)/k \cos(2\pi kx/T - \pi/2 - \psi r(k)), \qquad \text{(F 5)}$$

for k odd integer and $0 < x < T$. L_0 is mean luminance, a amplitude and T period (equation (12) in text). $r(x)$ varies randomly over the range -0.5–0.5 ψ, which determines the range of randomization, was $\pi/4$, $\pi/2$, π and 2π for (*a*), (*b*), (*c*) and (*d*) respectively. Except for (*d*), where the phases of all harmonics except the fundamental were completely random, the edge remains, embedded in varying amounts of noise.

Figure 12. Four periods of a two-dimensional waveform whose luminance distribution $f(x, y)$ is given by

$$f(x, y) = L_0 + 4a/\pi \sum_{k=1}^{\infty} G(k)/k \cos\{2\pi[(|y|/2T - 1/2) + k(1/8 - |y|/4T) + kx/T]\}, \qquad \text{(F 6)}$$

for k odd integer, $0 < x < 4T$ and $-T < y < T$. L_0 is mean luminance, a amplitude and T the horizontal period (equation (13) in text). The value of the argument at the point of arrival phase congruence changes continuously from the middle to the top and bottom rows. In addition, there is a group phase advance, which shifts the points of phase congruence of each row. Close up, the organization of the pattern follow the points of phase congruence, which form a chevron pointing left. From a distance, where the higher harmonics are unresolvable, the organization is dictated by the average luminance, which points rightwards.

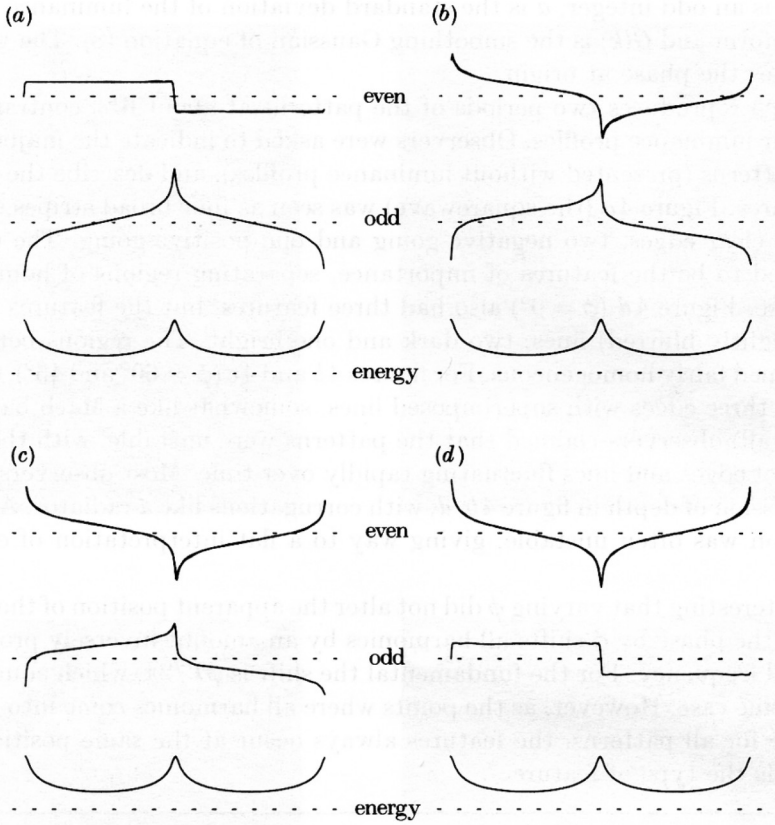


FIGURE 5. Sums of the convolutions of the waveforms of figure 4 with even-symmetric and odd-symmetric receptive fields of four scales (from figure 2). As the combined tuning of the receptive fields spans the whole range of the frequency content of the waveforms, the summed even-symmetric responses reproduce almost perfectly the original waveform and summed odd-symmetric responses of its Hilbert transform. Also shown are the local energy profiles, summed over all scales. Note that the local energy is identical for all the waveforms, and peaks at the visually salient feature. Whether the feature appears to be a line or an edge depends on the relative strengths of even to odd responses at those points.

Simulation

Figure 5 shows the sums of the convolutions of the waveforms of figure 4 with the even receptive fields of four scales (top profiles), and with odd receptive fields (middle profiles). As the receptive fields combine to form an almost flat filter, the summed even response reproduces almost perfectly the original waveforms, and the summed odd response its Hilbert transform. Also shown are the summed energy profiles, calculated by the method described in the Model section. The general form of the energy profile was similar at all scales, so we depict here only the sum. Despite the differences in the original waveforms, the local energy functions are all identical. They all peak at the points where the arrival phases of the harmonics are identical, irrespective of the value of the arrival phase at those points.

According to our model, local peaks in the energy function signal visually salient

features. These are interpreted as edges or bars (or both), depending on the strength of response from the odd and the even receptive fields at the points where local energy peaks. For the squarewave ($\phi = \pi/2$), the odd response is high at the peaks and the even response zero, indicating an edge. When $\phi = 0$, the even response is maximal and the odd response zero, indicating a line. At intermediate values of ϕ, both even and odd detectors respond, signalling both an edge and line of different relative strength. This predicted reasonably well the impressions of most observers.

What the model does not explain is why an impression of depth occurs with the intermediate values. The connection between apparent depth and phase is interesting, and bears further investigation. It may be connected with light distributions that occur with a three-dimensional object in natural lighting conditions.

Experiment 1

Some observers reported perceptual instability in the previous demonstration, for ϕ near $\pi/4$: the relative salience of the bars and edges fluctuated over time. The instability can be made more pronounced when the high spatial frequencies of the pattern are attenuated. Figure 6, plate 2, shows a waveform of squarewave amplitude and constant phase $\phi = \pi/4$, filtered with a low-pass Gaussian filter of standard deviation 2 (equation 8).

The Gaussian attenuates heavily the amplitude of all higher harmonics: the third harmonic by a factor of 0.36 (relative to the fundamental), the fifth by a factor of 0.05 and the seventh by a factor of 0.0025. The resultant waveform is similar to that created by Atkinson & Campbell (1974), who added only the first and third harmonics of a squarewave in various phases.

Readers can verify for themselves the perceptual instablity of figure 6. After a short period of viewing, its appearance begins to alternate rapidly between a 'squarewave' and a 'trianglewave' appearance, which can be more accurately described as a series of lines of alternating polarity. This perceptual alternation was described as 'monocular rivalry' by Atkinson & Campbell (1974).

Observers were asked to indicate the position of the centre (white) line when the pattern appeared as lines, and the centre edge when it appeared as an edge, by adjusting the position of a small superimposed spot. They were asked to wait until the apparent configuration was clearly one of lines or edges before making their judgement. Two observers participated in the task, and made ten measurements under each condition. The average settings of the two observers are indicated under the photograph of figure 6 (see also figure 7). The apparent position of the line and edge were remarkably similar, differing by about $0.0076\,T$ (where T is the period).

Several models of edge detection predict edges to occur at the points of maximal change of luminance (peaks in the first derivative, or zero-crossings in the second: see, for example, Marr & Hildreth (1980)). Lines should occur at the peak of the luminance profile, or possibly at the centroid of luminance mass above mean luminance (Watt & Morgan 1985). The positions of the maximal slope, and of the peaks and centroid of the luminance profiles, are indicated in figure 6. The average setting for the edge was $0.05\,T$ from the point of maximal slope. The average

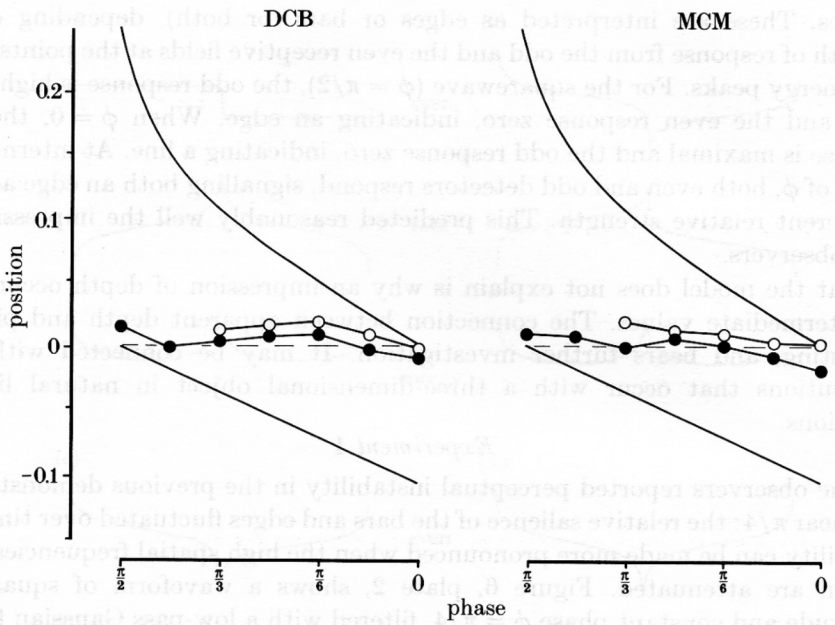


FIGURE 7. Average settings for two observers, DCB and MCM, for the apparent position of lines and edges in waveforms giving rise to monocular rivalry. All waveforms were low-pass filtered with the Gaussian of equation (8) with $\sigma = 2$ cycles per period. The phase spectra of the waveforms varied from $\pi/2$ to 0 (abscissa). Two cycles of waveform were displayed at spatial frequency 1 cycle deg^{-1} and contrast 0.01. Also shown are the positions of the peaks in luminance (upper curve), peak in energy profile (dashed line) and peak in luminance slope (lower curve). The apparent positions of lines and edges follow quite closely the position of the peak in local energy.

setting for the line was $0.07\,T$ from the luminance peak and $0.12\,T$ from the centroid. However, the peak in the energy profile (lower profile of figure 6) occurred $0.0021\,T$ from the perceived edge, and $0.0055\,T$ from the perceived line. That is to say, the local energy profile predicted the position of perceived lines and edges with considerable accuracy, whereas predictions of other models failed.

The apparent positions of lines and edges were then measured as a function of phase ϕ. As Atkinson and Campbell observed, rivalry is strongest and most rapid at phases around 45°, but most phases (other than those near 90°) produced some rivalry. Figure 7 shows the results for seven different phases for two observers. The apparent positions of the lines and edges was fairly constant under all conditions, and close to the peak in the local energy function (indicated by the dashed lines). There was a slight tendency for lines to appear to the right of edges, but the difference was seldom more than $0.01\,T$. Both lines and edges were seen near the peak in the local energy profile. Also shown on the graphs are the peaks and points of maximal slope of the waveforms. Neither of these is an adequate predictor of the perceived position of features.

Demonstration 2

The purpose of this demonstration was to investigate the effect of manipulating the amplitude spectra of waveforms. The phase spectra of the waveforms used

here were identical to those of demonstration 1, but the amplitudes of all harmonics were equal (discounting the smoothing Gaussian). The amplitude spectrum is that of a series of delta functions, rather than of a squarewave. The general equation for the waveforms is

$$f(x) = L_0 + 2a/\pi \sum_{k=1}^{\infty} G(k) \cos(2\pi kx/T - \phi), \tag{11}$$

where the sum is over odd values of k. The other symbols are as in equation (10).

Figure 8, plate 3, reproduces patterns for $\phi = \pi/2, \pi/3, \pi/4$ and 0 (90°, 60°, 45° and 0°), at about 10% contrast, together with their luminance profiles. As with the previous demonstration, observers reported three major features for each pattern, all in the same position. In figure 8*a* the features were largely edges (particularly at low contrast), in figure 8*d* lines, and the other figures gave some impression of both lines and edges (with lines dominating). As with demonstration 1, some observers reported a sensation of depth.

Figure 8*a* ($\phi = \pi/2$) is an interesting case. It appears to be a squarewave with alternately light and dark panels, particularly at low contrasts. However, as the luminance profile shows, the apparent panels are of almost equal luminance, except at the borders. Indeed the waveform, which can be considered to be a high-pass version of a squarewave, is similar in shape to that which creates the well known Craik–O'Brien–Cornsweet illusion (Craik 1940; O'Brien 1958; Cornsweet 1970), most powerful at low contrasts (Burr 1987).

The observations with this pattern suggest that the amplitude spectrum is not crucial. The general impression of these patterns is very similar to that of the patterns of figure 4, even though the amplitude spectrum (and hence the luminance profiles) were very different.

Simulation

Figure 9 shows the local energy functions, and summed even and odd responses for the waveforms of figure 8. The local energy functions are very similar to those of figure 5, as are the phase spectra (and hence the points of congruence of arrival phase) are the same. The only difference is that the high spatial frequencies are emphasized, so the smaller receptive fields make a greater contribution. Similarly, the response of odd and even fields at the maxima of the local energy functions is like that of figure 5.

It is interesting that the high responses of odd-symmetric receptive fields of figure 9*a* give rise to the perception of an edge, and of a brightness change, even though there is no accompanying change in luminance. This suggests that the signal for an edge is also the signal for a brightness change (even if it is not accompanied by a corresponding change in luminance), and may be the explanation for the Craik–O'Brien illusion (see Burr 1987).

Demonstration 3

The previous two demonstrations showed the effect of adding a constant to the phase spectrum of a waveform, leaving the spectrum flat. In this demonstration

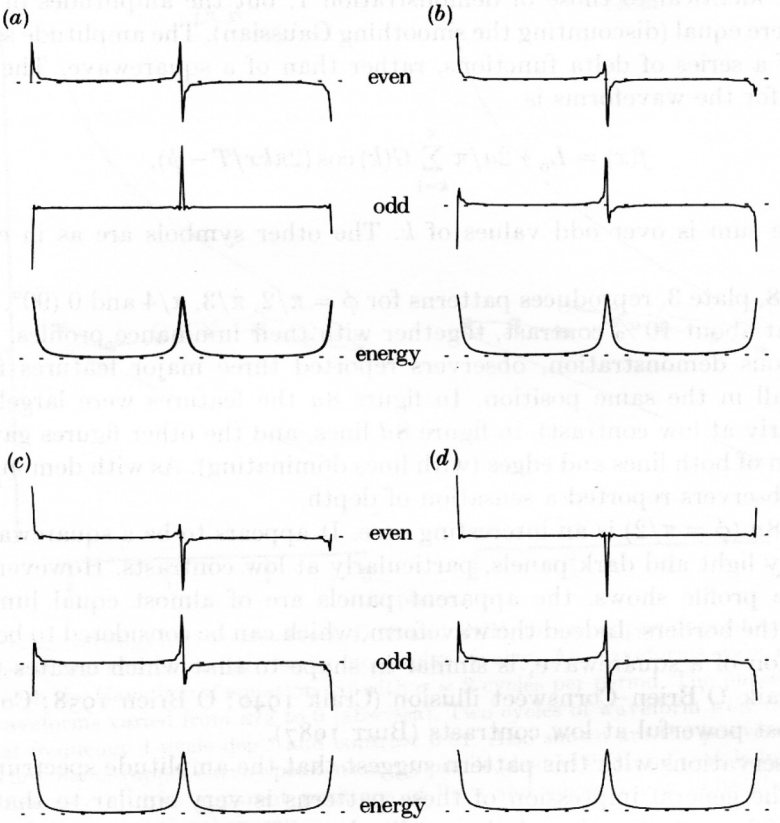


FIGURE 9. Local energy profiles of the waveforms of figure 8, summed over all scales, together with the summed response of even-symmetric detectors and odd-symmetric detectors. Again, the local energy is identical for all waveforms, and peaks at the visually salient feature. A clear edge is seen in figure 8*a*, although it is not accompanied by a luminance change, except near the borders.

the phases of all harmonics are perturbed randomly by varying degrees. The general formula for the waveforms is

$$f(x) = L_0 + 4a/\pi \sum_{k=1}^{\infty} G(k)/k \cos[2\pi kx/T - \pi/2 - \psi r(k)] \qquad (12)$$

where $r(k)$ varies randomly over the range -0.5–0.5, with $r(1) = 0$ (leaving the phase of the fundamental unaltered). The average phase for all the higher harmonics is $\pi/2$, varying randomly over a range ψ. The amount of variation will depend on the particular draw of random numbers on that occasion. Several waveforms were generated with independent random number sequences. Those reproduced in figure 10, plate 4, are fairly typical samples.

In figure 10 the value of ψ (range of perturbation of the phase spectrum) was $\pi/4$, $\pi/2$, π and 2π (45°, 90°, 180° and 360°). Observers all reported that the general impression of all patterns except figure 10*d* was of a squarewave, with varying amounts of added noise. Figure 10*d* appeared like random noise.

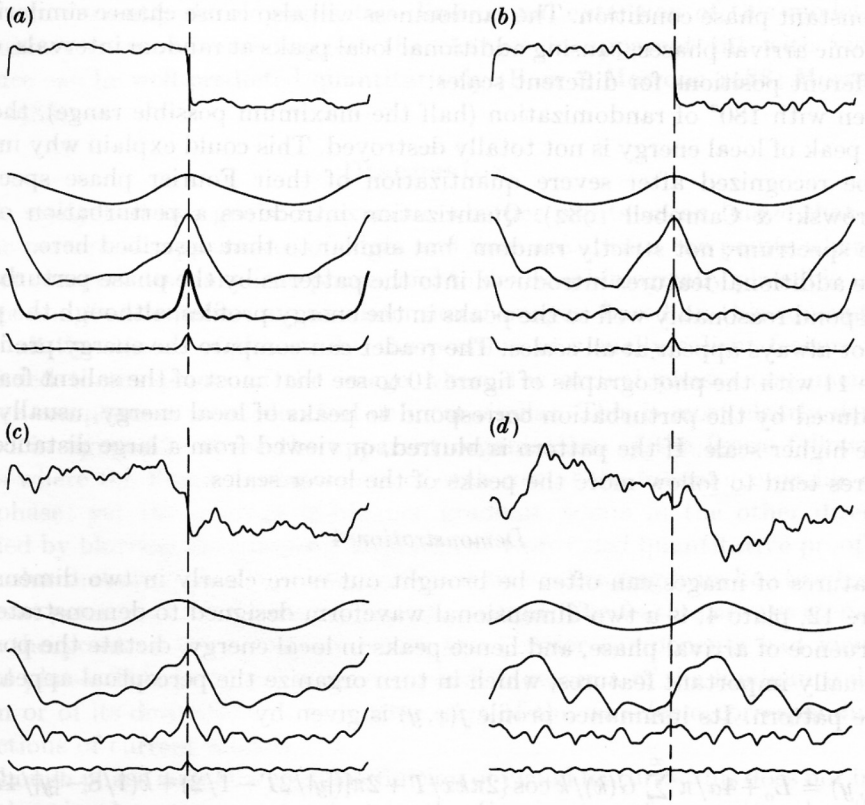


FIGURE 11. Profiles of the waveforms of figure 10, together with local energy profiles at each of the four scales. For all waveforms except (*d*), there are strong peaks at the centre and extremities of the energy functions at all scales, indicating the edge of the parent waveform. Other peaks introduced by phase perturbations do not necessarily match at all scales, but most correspond to perceived features in the patterns.

Simulation

Figure 11 shows the local energy profiles for each of the waveforms at each of the four separate scales. For the first three patterns, there are two strong peaks in energy at all scales, corresponding to the perceived edge in the patterns. Increasing the randomness factor, ψ, progressively decreased the amplitude of the central energy peak, and introduced additional peaks. However, even with 180° of random perturbation, the major peak occurred at the centre of the pattern. With 360° of perturbation (total randomness) the peak at this position disappeared completely.

The effect of phase randomization on the local energy profile can be understood by considering the relation between local energy amplitude and similarity of arrival phase (illustrated in figure 1). When the phase spectrum is not constant but partly randomized, it is improbable that all harmonics will have identical arrival phases at any point on the waveform. As the mean of the perturbation is zero, arrival phases will (on average) be most similar where $x = 0$ or $x = T/2$, but as the arrival phases are not identical, local energy at these points will be less than for

the constant phase condition. The randomness will also cause chance similarity of harmonic arrival phases, causing additional local peaks at random intervals, often at different positions for different scales.

Even with 180° of randomization (half the maximum possible range), the original peak of local energy is not totally destroyed. This could explain why images can be recognized after severe quantization of their Fourier phase spectrum (Piotrowski & Campbell 1982). Quantization introduces a perturbation of the phase spectrum; not strictly random, but similar to that described here.

The additional features introduced into the patterns by the phase perturbation correspond reasonably well to the peaks in the energy profiles, although the peaks did not always appear at all scales. The reader can compare the energy profiles of figure 11 with the photographs of figure 10 to see that most of the salient features introduced by the perturbation correspond to peaks of local energy, usually that of the higher scale. If the pattern is blurred, or viewed from a large distance, the features tend to follow more the peaks of the lower scales.

Demonstration 4

Features of images can often be brought out more clearly in two dimensions. Figure 12, plate 4, is a two-dimensional waveform designed to demonstrate how congruence of arrival phase, and hence peaks in local energy, dictate the position of visually important features, which in turn organize the perceptual appearance of the pattern. Its luminance profile $f(x, y)$ is given by

$$f(x,y) = L_0 + 4a/\pi \sum_{k=1}^{\infty} G(k)/k \cos\{2\pi kx/T + 2\pi[(|y|/2T - 1/2) + k(1/8 - |y|/4T)]\} \tag{13}$$

for k as an odd integer and $-T < y < T$.

Close to, the pattern appears as a chevron pointing left. However, this perceptual organization is actually contrary to that given by the average luminance distribution. To verify this, view the pattern from a large distance (or screw up your eyes): the chevron appears to point right.

The model proposed here has so far been developed only for one-dimensional patterns. However, some idea of how the model performs can be gained by considering the pattern row by row. For each row the arrival phases of all harmonics are identical at eight points. The points where the harmonics come into phase shift systematically rightward from the middle row outwards, because of the group phase advance (given by the term $k(\frac{1}{8} - |y|/4T)$). The peaks of local energy will follow the points of phase congruence, and dictate the perceptual organization of the pattern. The term $|y|/2T - \frac{1}{2}$ changes the value of the arrival phase (or argument) at these points, but as local energy is independent of the argument value, this does not affect the organization of energy peaks. It does, however, affect the luminance distribution. When the pattern is blurred or viewed from a distance, the higher harmonics are unresolvable, so the peaks in local energy become shallow or disappear entirely. When this occurs, the orientation is dictated by the average luminance, which creates a chevron pointing in the other direction. A similar effect has been observed with filtered checkerboards (Burr *et al.* 1986).

We are currently developing a two-dimensional extension of this model, and preliminary measurements suggest the changes in apparent tilt with viewing distance can be well predicted quantitatively (Burr & Morrone 1988; Morrone & Burr 1988).

Discussion

This study had two goals: to examine whether points on an image that have unique perceptual appearance of edges and lines are always the points where the Fourier components come into phase; and to devise a biologically plausible model of edge and line detection in one-dimensional images, based on this concept.

For all the stimuli of this study, the position of visually salient features corresponded to the points of the image where the arrival phases (argument) of Fourier components were identical or most similar. This is most clearly demonstrated in figure 12, where the apparent organization of the figure follows the points where the Fourier components of each row of the luminance profile come into phase; yet the average luminance gradient points in the other direction (verified by blurring the pattern). Experiment 1 provided quantitative proof. For a stimulus that alternated in appearance over time, changing from line to edge, the apparent position of the feature remained constant. The feature always seemed to be at the point of the waveform where the Fourier components had identical arrival phase. This point did not correspond to maxima of the luminance distribution or of its derivative, or any other part of the waveform corresponding to predictions of current models.

Although phase may be a useful definition of visual features, this does not imply that the visual system must calculate the Fourier transform, either locally or globally. Points where the Fourier components come into phase can be readily detected by calculating the local energy of the input waveform (using operators similar to those known to exist in the visual system) and searching for peaks. As proved in the Model section, peaks in local energy occur where the arrival phases of Fourier components are most similar.

The local energy model requires a basis of at least two operators, preferably orthogonal. Our implementation uses detectors with even- and odd-symmetric receptive fields, whose output is related via the Hilbert transform. We choose these filters partly on the basis of the existing (but scanty) psychophysical evidence, and partly to simplify the presentation of the model. Any orthogonal pair of filters of equal amplitude spectrum would do equally well. However, it is important to understand that the operation cannot be achieved with only one operator (by introducing, for example, a spatial displacement). Local energy is the vector length in a two-dimensional space, and cannot be calculated from information about only one dimension.

An intuitive understanding of local energy in the spatial domain is difficult. For simple patterns, such as isolated lines and edges, peaks in local energy correspond to peaks and zero-crossings of the image and its Hilbert transform. At these points the stimuli are locally symmetric. However, for complex patterns such as Mach bands and combinations of lines and edges (such as the demonstrations of this study), there are no such obvious parallels. The patterns with constant phase

spectra other than 0° or 90° (which can occur naturally, in conditions of oblique lighting) are not symmetric about the peaks of local energy (at any scale), and neither the even- nor odd-symmetric fields have maximal or zero response. For this reason, we assert that coincidence of arrival phase of Fourier components is a useful and valid definition of image features.

Some readers may be surprised that an energy model should be phase sensitive. In their recent model of motion perception, Adelson & Bergen (1985) suggested a similar operator for motion detectors, because it provides a 'phase-independent' output. Klein & Levi (1985) also use local energy (which they term the 'pythagorean sum') 'to preserve full information about the magnitude of stimulation, while discarding phase information' (page 1179). For pure harmonic signals (sine waves) the transformation is indeed phase independent, as $\sin^2(x)+\cos^2(x) = 1$: the energy profile will be flat. For complex stimuli, however, the local energy profile depends critically on the relations between the phases of the various harmonics.

The role of spatial phase

The present results help reconcile some seemingly paradoxical observations of visual phase perception: on the one hand, the phase spectrum contains most of the image-related information of natural images (see, for example, Openheim & Lim 1981; Piotrowski & Campbell 1982); on the other hand, phase discrimination is poor (Burr 1980; Badcock 1984*a*, *b*), 'phase perception' unstable (Atkinson & Campbell 1974) and phase can survive severe quantization (Piotrowski & Campbell 1982). The Fourier phase spectrum is crucial (more so than the amplitude spectrum) because the phase spectrum determines the points where the arrival phases of harmonics will be coincident (or most similar), and hence the position and nature of visually salient features. However, there is considerable room for random perturbations (of the type introduced by quantization) before the grouping of arrival phases congruency is completely destroyed.

Demonstration 2 showed that amplitude information was less important than phase for perception. Provided that the arrival phases of all the harmonics were all aligned at 90° (sine) phase, the impression was of a squarewave, with alternately bright and dark bars (even though they had the same luminance).

The model can readily accommodate the visual instability produced by certain phase relationships (monocular rivalry). Strongest and most rapid rivalry occurs when the arrival phases of all harmonics are around 45° (Atkinson & Campbell 1974). At this point, both odd- and even-symmetric detectors respond strongly. Should there be random variations in the relative strength of response over time, one would expect the impression of edge and line to fluctuate over time. Some support for this idea is given by the fact that monocular rivalry is strongest at low contrasts, and with low-pass filtered images (to which few detectors respond), where the random fluctuations should be greatest.

Under conditions of monocular rivalry, the model successfully predicted the apparent positions of the features. Although the nature of the feature fluctuated over time, changing from bar to edge, its position remained constant. Both features were seen where the Fourier components came into phase, where local energy was maximal.

Alternative models

Most models of feature analysis rely heavily on linear filter operators. For example, the seminal model of Marr and his colleagues (Marr & Hildreth 1980; Marr 1982) employs the Laplacian-of-Gaussian operator to filter the image at different scales. Coincidence of zero-crossings across scales provides the basis for a schematic description of the image (the 'primal sketch').

For simple edges, such as those of figure 3*a*, coincidence of zero-crossings occurs at the edges. For lines (such as figure 3*d*), there is no coincidence in zero-crossings, so no edges are marked. Thus the model cannot detect lines, even though they are salient visual features. The model encounters greatest difficulties with features that are combinations of edges and lines, such as those with constant phase of 45° (figure 3*b*). Again, zero-crossings do not coincide across scales, but edges are seen by observers. This limitation is most evident with the low-pass filtered waveforms of experiment 1. There the zero-crossings (at all scales) occur at the points of maximal slope (where the second derivative crosses zero); but edges were not perceived at that point. Furthermore, neither the original Marr & Hildreth (1980) model nor subsequent developments (see, for example, Yuille & Poggio 1985) predicts the perceptual alternations between lines and edges. In another publication, we have demonstrated that zero-crossings do not predict Mach bands, whereas our model does so with precision (Ross *et al.* 1988).

Other, more recent, models proceed along similar lines to that of Marr. For example, Watt & Morgan's (1985) 'MIRAGE' filters the image at various scales, half-wave rectifies at each scale, then searches for the centroid of each zero-bounded response after summing separately all positive and negative responses. With simple patterns such as squarewaves, the approach is reasonably successful. However, it fails with the waveform of experiment 1. After filtering, half-wave rectification and summing across scales, the resultant output is similar to a full-wave rectified version of the original. According to Watt and Morgan's rules of interpretation, the visual system should see a series of wide bars (squarewave) centred at the centroid of each zero-bounded mass and extending to the point where the waveform crosses mean luminance. Edges should appear at the mean-luminance cross (which is also the point of maximal slope), and bars at the centroids. These predictions are not borne out by the data.

Another difficulty for both the above models is that the spatial filters introduce ripples, more or less pronounced depending on the bandwidth of the filter. The ripples cause spurious peaks or zero-crossings, which can lead to marking features where there are none. To discard the spurious signals, both models require 'rules' for interpreting the output. Marr searches for coincidence in zero-crossings across scales, which eliminates the ripples (as they occur at different points at different scales), but also eliminates many real features, such as lines. Watt and Morgan evaluate the response sequence after filtering and rectification, and discriminate between edges, lines and ripples primarily on the basis of the symmetry of the output. Again, the approach has only limited success, particularly if two features are nearby (or superimposed). For our model, neither approach is necessary. For most stimuli, the local energy function is smooth. The ripples introduced by the

filters are automatically eliminated when the outputs of the complementary filters are combined. Thus the problem of spurious peaks or zero-crossings, and the need to devise arbitrary cognitive rules (which may be difficult to implement rapidly and effortlessly with biological hardware) do not arise.

There is also no need for cognitive rules to distinguish edges from lines. That information is contained in the linear stage of the model. If at the point of local energy maxima there is a response from odd-symmetric receptive fields, an edge is seen. If even-symmetric receptive fields respond there, a line is seen. If both respond, both edges and lines are seen (alternating in relative strength under some conditions).

Role of spatial filters

Our model commences with four linear filters, of approximately 1.5 octave bandwidth (chosen because of the experimental evidence for these filters in human and animal vision). However, the particular bandwidths and shape of the filters are not crucial for the model: nor indeed does the model require multiple filters. We have shown elsewhere that a single pair of matched filters of broad bandwidth can detect and locate image features in natural images, and may be a useful tool for artificial visual systems, where computational speed is important (Morrone & Owens 1987).

With the proposed model, spatial filtering of the image into narrow bands is not the crucial operation for feature detection (cf. Pollen *et al.* 1971; Marr & Hildreth 1982; Robson 1980). The primary function of the filters is to create two functions of identical amplitude spectrum, differing in phase spectra by 90° (Hilbert transforms of each other). Small variations in either the amplitude or the phase response can lead to distortions in the energy profiles, giving rise to false peaks. Variations in both the amplitude and the phase response can be minimized by restricting the bandwidth of the filters. Indeed, most broad-band electronic phases filters are 'multipole', comprising many narrow band filters.

Apart from providing robust filters of constant phase, the limited bandwidth of visual detectors probably serves various other functions, including reduction of image noise (see, for example, Georgeson & Sullivan 1975; Sakitt & Barlow 1982; Watt & Morgan 1985), increased encoding and transmission efficiency (see, for example, Burt & Adelson 1983), and possibly providing the visual system with a convenient mechanism for adjusting its resolution to the prevailing noise levels, which vary dramatically with luminance (Ross & Campbell 1978).

At present there is little evidence to suggest how the output at various scales may be combined. For this study, we assumed that energy is calculated at each scale, and that the local maxima at each scale all signal a feature. This process predicted reasonably well all the features of the complex waveforms of figure 10, where there was not always correspondence across scales. However, further experimentation is in progress to determine the rules of combining information from the filters of various frequency preferences in situations where they do not correspond across scales.

Another obvious limitation of the model in its present form is that it applies only to one-dimensional stimuli. In extending the model to two dimensions there

is a major restriction: the Hilbert transform is a one-dimensional transform. To generate two-dimensional filters in quadrature phase it is necessary to introduce an orientation bias in the filters. For two-dimensional images local energy must be computed separately for several different orientations, and the output combined in some way. This is not an unreasonable limitation, as the major property of cortical neurones is their orientation selectivity (Hubel & Wiesel 1962, 1977). A preliminary report of the two-dimensional extension to the model shows how local energy calculated from oriented matched filters can detect and locate features in two-dimensional images with reasonable accuracy (Burr & Morrone 1988; Morrone & Burr 1988).

Physiological mechanisms

There exist two classes of cells in the visual system: quasi-linear cells, such as X-retinal and -geniculate cells and simple cortical cells, and inherently nonlinear cells, such as Y-retinal and -geniculate cells and complex cortical cells. The role of the two classes of cells has been the subject of much speculation, but is still unclear (for review see Lennie (1980)).

Both linear and nonlinear operations are essential for our model. The linear operation requires pairs of filters, matched in amplitude response and differing in phase response. The receptive fields of these filters resemble closely those described by Hubel & Wiesel (1962, 1977), and subsequent research (see, for example, Maffei *et al.* 1979; Kulikowski & Bishop 1983). The phase response of adjacent simple cells tends to differ by 90° (Pollen & Ronner 1981), making them ideal candidates. Field & Tolhurst (1986) suggest that the phase preference of simple cells may not group around 0° and 90°, but vary over the entire range with equal frequency. This finding, if confirmed, would not change significantly the model presented. To extract energy, any pair of filters of the same orientation preference and in quadrature phase will suffice (see, for example, Gabor 1946).

The nonlinear stage of the model involves squaring (a second-order nonlinearity) and summing the output of the matched filters. Complex cortical cells are ideal candidates for this task. They modulate to the second harmonic (and other even harmonics) to counter-phased gratings (Movshon *et al.* 1979*b*; Spitzer & Hochstein 1985) indicating a second-order nonlinearity, and their response to drifting sinusoidal gratings is uniform, reflecting the uniform energy in sine-wave gratings. To test this idea, we intend to measure the spike activity of complex cells in response to trapezoids, and waveforms such as those used in this study.

If this line of approach proves successful, it would suggest separate roles for the linear and nonlinear neural elements. The nonlinear cells are responsible for the *detection* and *location* of features (by calculating energy) and the linear cells for the *identification* of features (lines or edges).

We thank John Ross, Horace Barlow and John Robson for discussions and comments on the manuscript. M.C.M. was on leave from the Scuola Normale Superiore, Pisa, Italy, and received a Queen Elizabeth II Fellowship from the Australian Department of Science. D.B. is an NH and MRC Senior Research Fellow.

References

Adelson, E. H. & Bergen, J. R. 1985 Spatio-temporal energy models for the perception of motion. *J. opt. Soc. Am.* A **2**, 284–299.

Anderson, S. J. & Burr, D. C. 1985 Spatial and temporal selectivity of the human motion detection system. *Vision Res.* **25**, 1147–1154.

Anderson, S. J. & Burr, D. C. 1987 Receptive field sizes of human motion detectors. *Vision Res.* **27**, 621–635.

Atkinson, J. & Campbell, F. W. 1974 The effect of phase on the perception of compound gratings. *Vision Res.* **14**, 159–162.

Badcock, D. R. 1984*a* How do we discriminate spatial phase? *Vision Res.* **24**, 1847–1857.

Badcock, D. R. 1984*b* Spatial phase or luminance profile discrimination? *Vision Res.* **24**, 613–623.

Barlow, H. B. 1959 Possible principles underlying the transformations of sensory messages. In *Sensory communication* (ed. W. A. Rosenblith), pp. 217–234. Cambridge, Massachusetts: MIT Press.

Blakemore, C. & Campbell, F. W. 1969 On the existence of neurones in the visual system selectively sensitive to the orientation and size of retinal images. *J. Physiol., Lond.* **225**, 437–455.

Burr, D. C. 1980 Sensitivity to spatial phase. *Vision Res.* **20**, 391–396.

Burr, D. C. 1987 Implications of the Craik–O'Brien illusion for brightness perception. *Vision Res.* **27**, 1903–1913.

Burr, D. C. & Morrone, M. C. 1988 Edge detection in biological and artificial visual systems. In *Vision: coding and efficiency* (ed. C. Blakemore). Cambridge University Press. (In the press.)

Burr, D. C., Morrone, M. C. & Ross, J. 1986 Local and global visual analysis. *Vision Res.* **26**, 749–757.

Burr, D. C., Morrone, M. C. & Spinelli, D. 1988 Evidence for edge and bar detectors in human vision. *Vision Res.* (In the press.)

Burt, P. J. & Adelson, E. H. 1983 The Laplacian pyramid as a compact image code. *IEEE Trans. Comm.* **31**, 532–540.

Canny, J. F. 1983 *Finding edges and lines in images.* MIT Artificial Intelligence Laboratory Technical Report no. 720.

Cornsweet, T. N. 1970 *Visual perception.* New York: Academic Press.

Craik, K. J. W. 1940 *Visual adaptation.* Ph.D. thesis, University of Cambridge.

Field, D. J. & Nachmias, J. 1984 Phase reversal discrimination. *Vision Res.* **24**, 333–340.

Field, D. J. & Tolhurst, D. J. 1986 The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proc. R. Soc. Lond.* B **226**, 379–399.

Gabor, D. 1946 Theory of communication. *IEEE J.* **93**, 429–457.

Garnett, J. B. 1981 *Bounded analytic functions.* New York: Academic Press.

Georgeson, M. A. & Sullivan, G. D. 1975 Contrast constancy: deblurring in human vision by spatial frequency channels. *J. Physiol., Lond.* **252**, 627–656.

Hewitt, E. & Hewitt, R. E. 1979 The Gibbs–Wilbraham phenomenon: an episode in Fourier analysis. *Archs Hist. exact Sci.* **21**, 130–160.

Hubel, D. H. & Wiesel, T. N. 1962 Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol., Lond.* **160**, 106–154.

Hubel, D. H. & Wiesel, T. N. 1977 Architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond.* B **198**, 1–59.

Klein, S. A. & Levi, D. M. 1985 Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation. *J. opt. Soc. Am.* A **2**, 1170–1190.

Kulikowski, J. J. & Bishop, P. O. 1981 Linear analysis of the responses of simple cells in the cat visual cortex. *Exp. Brain Res.* **44**, 386–400.

Kulikowski, J. J. & King-Smith, P. E. 1973 Spatial arrangements of line, edges and grating detectors revealed by subthreshold summation. *Vision Res.* **13**, 1455–1478.

Landy, M. S., Cohen, Y. & Sperling, G. 1984*a* HIPS: a UNIX-based image processing system. *Comput. Vision, Graphics Image Processing* **25**, 331–347.

Landy, M. S., Cohen, Y. & Sperling, G. 1984*b* HIPS: image processing under UNIX. Software and applications. *Behav. Res. Meth.* **16**, 199–216.

Lennie, P. 1980 Parallel visual pathways: a review. *Vision Res.* **20**, 561–594.
Mach, E. 1865 Über die Wirkung der raumlichen Vertheilung des Lichreizes auf die Neztzhaut. *Sber. Akad. Wiss. Wien* **54**, 303–322.
Maffei, L. & Fiorentini, A. 1973 The visual cortex as a spatial frequency analyzer. *Vision Res.* **13**, 1255–1267.
Maffei, L., Morrone, M. C., Pirchio, M. & Sandini, G. 1979 Response of visual cortical cells to periodic and non-periodic stimuli. *J. Physiol., Lond.* **296**, 27–47.
Marr, D. 1976 Early processing of visual information. *Phil. Trans. R. Soc. Lond.* B **275**, 485–526.
Marr, D. 1982 *Vision*. San Francisco: Freeman.
Marr, D. & Hildreth, E. 1980 Theory of edge detection. *Proc. R. Soc. Lond.* B **207**, 187–217.
Morrone, M. C. & Burr, D. C. 1988 A model of feature detection based on matched filters. *J. opt. Soc. Am.* (Submitted.)
Morrone, M. C. & Owens, R. 1987 Edge detection by local energy. *Pattern recognition Letts.* **6**, 303–313.
Morrone, M. C., Ross, J., Burr, D. C. & Owens, R. 1986 Mach bands depend on spatial phase. *Nature, Lond.* **324**, 250–253.
Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. 1978*a* Receptive field organization of complex cells in the cat's striate cortex. *J. Physiol., Lond.* **283**, 79–99.
Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. 1978*b* Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol., Lond.* **283**, 53–77.
O'Brien, V. 1958 Contour perception, illusion and reality. *J. opt. Soc. Am.* **48**, 112–119.
Openheim, A. V. & Lim, J. S. 1981 The importance of phase in signals. *Proc. Inst. elect. Electron. Engrs* **69**, 529–541.
Piotrowski, L. N. & Campbell, F. W. 1982 A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception* **11**, 337–346.
Pollen, D. A., Lee, J. R. & Taylor, J. H. 1971 How does the striate cortex begin the reconstruction of the visual world? *Science, Wash.* **173**, 74–77.
Pollen, D. A. & Ronner, S. F. 1981 Phase relationships between adjacent simple cells in the visual cortex. *Science, Wash.* **212**, 1409–1411.
Robson, J. G. 1980 Neural images: The physiological basis of spatial vision. In *Visual coding and adaptability* (ed. L. S. Harris, L. Erlbaum *et al.*), pp. 177–214. New Jersey: Hillsborough.
Ross, J. & Campbell, F. W. 1978 Why we do not see photons. *Nature, Lond.* **275**, 541–542.
Ross, J., Morrone, M. C. & Burr, D. C. 1988 The conditions under which Mach bands are visible. *Vision Res.* (In the press.)
Sakitt, B. & Barlow, H. B. 1982 A model for the economical encoding of the visual image in cerebral cortex. *Biol. Cybernet.* **43**, 97–108.
Schiller, P. H., Finlay, B. L. & Volman, S. 1976 Quantitative studies of single cells properties in monkey striate cortex. III. Spatial frequency. *J. Neurophysiol.* **39**, 1334–1351.
Shapley, R. M. & Tolhurst, D. J. 1973 Edge detectors in human vision. *J. Physiol., Lond.* **229**, 165–183.
Spitzer, H. & Hochstein, S. 1985 A complex-cell receptive field model. *J. Neurophysiol.* **53**, 1266–1286.
Tolhurst, D. J. 1972 On the possible existence of edge-detector neurones in the human visual cortex. *Vision Res.* **12**, 797–804.
Watt, R. J. & Morgan, M. J. 1985 A theory of the primitive spatial code in human vision. *Vision Res.* **25**, 1661–1674.
Wilson, H. R., McFarlane, D. R. & Phillips, G. C. 1983 Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Res.* **23**, 873–882.
Yuille, A. L. & Poggio, T. 1985 Fingerprint theorems for zero crossings. *J. opt. Soc. Am.* A **2**, 683–692.